Attorney Docket No. 349560

THE UNITED STATES PATENT AND TRADEMARK 🕅

0349580010-01060

n 1892Appincation of:

HANDCARRIED

MARC ALIZON et al.

Serial No.: 158,652

Group Art Unit: 18/05

Filed: February 22, 1988

Examiner: J. Railey

For: CLONED DNA SEQUENCES RELATED TO THE GENOMIC RNA

OF LYMPHADENOPATHY ASSOCIATED VIRUS (LAV) AND

PROTEINS ENCODED BY SAID LAV GENOMIC RNA

Honorable Commissioner of Patents and Trademarks Washington, D.C. 20231

Sir:

SUBMISSION OF COPY OF PRIORITY DOCUMENT

Further to the Examiner's telephone request, enclosed herewith is a copy of United Kingdom Priority Document No. 8429099.

Respectfully submitted,

FINNEGAN, HENDERSON, FARABOW, GARRETT & DUNNER

Rv:

Kenneth J. Meyers

Reg. No. 25,146

Dated: February 6, 1992

LAW OFFICES
FINNEGAN, HENDERSON
FARABOW, GARRETT
& DUNNER
1300 I STREET, N. W.
WASHINGTON, DC 20005
1-202-408-4000



19/11/84 B3662 PAT*** 10.00

The Comptroller
The Patent Office All ROOM
25 Southamptor Balldings
London, WC24 1AY FEB
15

Fee: E10.00

1984 2 9 0 9 9

600 12112 FEB 10 1992

8429 SABERINE

THE APPL	GRANT OF A PATENT IS	REQUESTED BY TO	HE UNDERSIGNED ON THE	BASIS OF TH	PRESENT
ī	Agent's Reference	JJD/EAF/26804			Vermen
11	Title of Invention	LYMPHADENOPAT	COUENCES RELATED TO THE ASSOCIATED VIRUS ALD LAV GENOMIC RNA.	RE GENOMIC (LAV) AND	RNA OF PROTEINS
111	Applicant or Applic				
	Name (First or only		INSTITUT PASTEUR		
	Country FR	ANCE	State	AC	P Code No
	Address25-20	Paris Codov 15	E.	••••••••	***************************************
			, France.		
			an one) Centre Natio		
	de la recherche	SCIENCIFICIE	Country FRANCE	St	ate
			<u>ce</u>		
	73007	raits, France.	***************************************	••••••	
IV	Inventor (see note 3)	Peranchica suggests Cast		D020000
			or (b) A statement on Patents	Form No. 7/7	7 is/will be furnished
<u>v</u>	Name of Agent (if a	ny) (See note 4)	Reddie & Gro	se	ADP CODE NO
۷١	Address for Service	(See note 5)	16 Theobalds Road		
			London WCIX 8PL		
VII					
411	Declaration of Prior	ity (See note 6)			
711	Declaration of Priori Country		g date	File nun	nber
********			g date	File nun	nber
********			g date		nber
************			g date	25	nber
***************************************			g date	25	nber
VIII	Country	Filing		25 SOUE .	
*******	Country	Filing	under Section 8(3), 12(6),	25 SOUE 'AIR TO 14 AIR TO 14 AIR TO 15 AIR TO	A) (SA note 7)

NOTES:

- This form, when completed, should be brought or sent to the Patent Office together with the prescribed fee and two copies of the description of the invention, and of any drawings.
- 2. Enter the name and address of each applicant. Names of individuals should be indicated in full and the surname or family name should be underlined. The names of all partners in a firm must be given in full. Bodies corporate should be designated by their corporate name and the country of incorporation and, where appropriate, the state of incorporation within that country should be entered where provided. Full corporate details, eg "a corporation organised and existing under the laws of the State of Delaware, United States of America," trading styles, eg "trading as xyz company", nationality, and former names, eg "formerly [known as] ABC Ltd." are not required and should not be given. Also enter applicant(s) ADP Code No. (if known).
- 3. Where the applicant or applicants is/are the sole inventor or the joint inventors, the declaration (a) to that effect at IV should be completed, and the alternative statement (b) deleted. If, however, this is not the case the declaration (a) should be struct: out and a statement will then be required to be filed upon Patent Form No 7/77.
- If the applicant has appointed an agent to act on his behalf, the agent's name and the address of his place of business should be indicated in the spaces available at V and VI. Also insert agent's ADP Code No. (if known) in the box provided.
- An address for service in the United Kingdom to which all documents may be sent must be stated at VI. It is recommended that a telephone number be provided if an agent is not appointed.
- The declaration of priority at VII should state the date of the previous filing and the country in which it was made and indicate the file number, if available.
- When an application is made by virtue of section 8(3), 12(6), 15(4), or 37(4) the appropriate section should be identified at VIII and the number of the earlier application or any patent granted thereon identified.
- Attention is directed to rules 90 and 106 of the Patent Rules 1982.
- 9. Attention of applicants is drawn to the desirability of avoiding publication of inventions relating to any article, material or device intended or adapted for use in war (Official Secrets Acts, 1911 and 1920). In addition after an application for a patent has been filed at the Patent Office the comptroller will consider whether publication or communication of the invention should be prohibited or restricted under section 22 of the Act and will inform the applicant if such prohibition is necessary.
- 10. Applicants resident in the United Kingdom are also reminded that, under the provisions of section 23 applications may not be filled abroad without written permission or unless an application has been filled not less than six weeks previously in the United Kingdom for a patent for the same invention and no direction prohibiting publication or communication has been given or any such direction has been received.

1

The invention relates to cloned DNA sequences indistinguishable from genomic RNA and DNA of lymphadenopathy-associated virus (LAV), a process for their preparation and their uses. It relates more particularly to stable probes including a DNA sequence which can be used for the detection of the LAV virus or related viruses or DNA proviruses in any medium, particularly biological samples containing any of them. The invention also relates to polypeptides, whether glycosylated or not, encoded by said DNA sequences.

10

15

20

30

35

Lymphadenopathy-associated virus (LAV) is a human retrovirus first isolated from the lymph node of a homosexual patient with lymphadenopathy syndrome, frequently a prodrome or a benign form of acquired immune deficiency syndrome (AIDS). Subsequently other LAV isolates have been recovered from patients with AIDS or pre-AIDS. All available data are consistent with the virus being the causative agent of AIDS.

A method for cloning such DNA sequences has already been disclosed in British Patent Application Nr.

after made to that application as concerns subject matter in common with the further improvements to the invention disclosed herein.

The present invention aims at providing additional new means which should not only also be useful for the detection of LAV or related viruses (hereafter more generally referred to as "LAV viruses"), but also have more versatility, particularly in detecting specific parts of the genomic DNA of said viruses whose expression products are not always directly detectable by immunological methods.

The present invention further sims at providing

polypeptides containing sequences in common with polypeptides encoded by the LAV genomic RNA. It relates even more particularly to polypeptides comprising antigenic determinants included in the proteins encoded and expressed by the LAV genome occuring in nature. An additional object of invention is to further provide means for the detection of proteins related to LAV virus, particularly for the diagnosis of AIDS or pre-AIDS or, to the contrary, for the detection of antibodies against the LAV virus or related therewith, particularly in patients proteins afflicted with AIDS or pre-AIDS or more generally in asymtomatic carriers and in blood-related products. Finally the invention also aims at providing immunogenic and polypeptides. more particularly protective polypeptides for use in the preparation of vaccine compositions against AIDS or related syndroms.

The present invention relates to additional DNA fragments, hybridizable with the genomic RNA of LAV as they will be disclosed hereafter, as well as with additional cDNA variants corresponding to the whole genomes of LAV viruses. It further relates to DNA recombinants containing said DNAs or cDNA fragments.

The invention relates more particularly to a cDNA variant corresponding to the whole of LAV retroviral genomes, which is characterized by a series of restriction sites in the order hereafter (from the 5° end to the 3° end).

The coordinates of the successive sites of the whole LAV genome (restriction map) are indicated hereafter too, with respect to the Hind III site (selected as of coordinate 1) which is located in the R region. The coordinates are estimated with an accuracy of 1 200 bp':

Hind III	· O
Sac I	50
Hind III	520
Pat I	800
Hind III	1 100

10

15

20

25

30

		•
	8gl II	1 500
	Kpn I	3 500
	Kpn I	3 900
	Eco RI	4 100
5	Eco RI	5 300
•	Sal I	5 500
	Kpn I	6 100
	Bgl II	8 500
	Bgl II	7 600
10	Hind III	7 850
	Bam HI	8 150
	Xho I	8 600
	Kpn I	8 700
	8g1 II	8 750
15	8gl II	9 150
	Sac I	9 200
	Hind III	9 250

20

25

30

Another DNA variant according to this invention optionally contains an additional Hind III approximately at the 5 550 coordinate.

Reference is further made to fig. 1 which shows a more detailed restriction map of said whole-DNA $(\lambda J19)$.

An even more detailed nucleotidie sequence of a preferred DNA according to the invention is shown in fig. 4-12 hereafter.

The invention further relates to other preferred DNA fragments which will be referred to hereafter.

Additional features of the invention will appear in the course of the non-limitative disclosure of additional features of preferred DNAs of the invention, as well as of preferred polypeptides according to the invention. Reference will further be had to the drawings in which:

-' fig. 1 is the restriction map of a complete LAV genome (clone $\lambda J19$);

35 - figs. 2 and 3 show diagrammatically parts of the three

possible reading phases of LAV genomic RNA, including the open reading frames (ORF) apparent in each of said reading phases;

- figs. 4-12 show the successive nucleotidic sequences of a complete LAV genome. The possible peptidie sequences in relation to the three possible reading phases related to the nucleotidie sequences shown are also indicated;
- figs. 13-18 reiterate the sequence of part of the LAV genome containing the genes coding for the enveloppe proteins, with particular boxed peptidic sequences which corresponds to groups which normally carry glycosyl groups.

The sequencing and determination of sites of particular interest was carried out on a phage recombinant corresponding to AJ19 disclosed in the abovesaid British Patent application Nr. 84 23659. A method for preparing it is disclosed in that application.

The whole recombinant phage DNA of clone \$\lambda{J19}\$ (disclosed in the earlier application) was sonicated according to the protocol of DEININGER (1983), Analytical Biochem. 129, 216. the DNA was repaired by a Klenow reaction for 12 hours at 16°C. The DNA was electrophoresed through 0.8 % agarose gel and DNA in the size range of 300-600 bp was cut out and electroeluted and precipitated. Resuspended DNA (in 10 mM Tris, pH 8; 0.1 mM EDTA) was ligated into M13mp8 RF DNA (cut by the restriction enzyme Smal and subsequently alkaline phosphated), using T4 DNA-and RNA-ligases (Manistis T et al (1982) - Molecular cloning - Cold Spring Harbor Laboratory). An £. Colistrain designated as TG1 was used for further study. This strain has the following genotype:

Alac pro, supE, thi.F'traD36, proA8, lacIq, ZAM15,r

This \underline{F} , \underline{coli} TGI strain has the peculiarity of enabling recombinants to be recognized easily. The blue colour of the cells transfected with plasmids which did

10

15

20

25

not recombine with a fragment of LAV DNA is not modified. To the contrary cells transfected by a recombinant plasmid containing a LAV DNA fragment yield white colonies. The technique which was used is disclosed in Gene (1983), 26, 101.

This strain was transformed with the ligation mix using the Hanahan method (Hanahan D (1983) J. Mol. Biol. 166, 557). Cells were plated out on tryptone-agarose plate with IPTG and X-gal in soft agarose. White plaques were either picked and screened or screened directly in situ using nitrocellulose filters. Their DNAs were hybridized with nick-translated DNA inserts of pUC18 Hind III subclones of λ J19, this permitted the isolation of the plasmids or subclones of A which are identified in the table hereafter. In relation to this table it should also be noted that the designation of each plasmid is followed by the deposition number of a cell culture of £. coli TGI containing the corresponding plasmid at the "Collection Nationale des Cultures de Micro-organismes" (C.N.C.M.) of the Pasteur Institute in Paris, France. A non-transformed TGI cell line was also deposited at the C.N.C.M. under Nr. I-364. All these deposits took place on November 15, 1984. The sizes of the corresponding inserts derived from the LAV genome have also been indicated.

25

10

15

Essential features of the recombinant plasmids

- pJ19 - 1 plasmid (1-365)0.5 kb

Hind III - Sac I - Hind III

- pJ19 - 17 plasmid (I-367)

10

Hind III - Pst 1 - Hind. III

- pJ19 - 6 plasmid (I-366) 1.5 kb

Hind III (5') 15

Bam HI

Xho I

Kpn I

8gl II

Sac I (3') 20

Hind III

- pJ19-13 plasmid (I-368) 8.7 kb

Hind III (5') 25

Bgl II

Kon I

Kpn I

Eco RI

30 Eco RI

Sal I

Kpn I

8g1 II

Bgl II

35 Hind III (3') Positively hybridizing M13 phage plates were grown up for 5 hours and the single-stranded DNAs were extracted.

M13mp8 subclones of AJ19 DNAs were sequenced according to the dideoxy method and technology devised by Sanger et al (Sanger et al (1977), Proc. Natl. Acad. Sci. USA, 74, 5483 and M13 cloning and sequencing handbook, AMERSHAM (1983). the 17-mer oligonucleotide primer a-35 SdATP (400Ci/mmol, AMERSHAM), and 0.5X-5X buffer gradient gels (Biggen M.D. et al (1983, Proc. Natl. Acad. Sci. USA, 50, 3983) were used. Gels were read and put into the computer under the programs of Staden (Staden R. (1982), Nucl. Acids Res. 10, 4731). All the appropriate references and methods can be found in the AMERSHAM M13 cloning and sequencing handbook.

The complete sequence of AJ19 was deduced from the experiments as further disclosed hereafter.

Figs. 4-12 provide the DNA nucleotidie sequence of the complete genome of LAV. The numbering of the nucleotides starts from a left most Hind III restriction site (5'AAG..) of the restriction map. The numbering occurs in tens whereby the last zero number of each of the numbers occuring on the drawings is located just below the nucleotide corresponding to the nucleotides designated. I.e. the nucleotide at position 10 is T, the nucleotide at position 20 is C, etc..

Above each of the lines of the successive nucleotidic sequences there are provided three lines of single letters corresponding to the aminoacid sequence deduced from the DNA sequence (using the genetic code) for each at the three reading phases, whereby said single letters have the following meanings.

A : alanine

10

15

20

25

30

35

R : arginine

K : lysine

H : histidine

C : cysteine

W : tryptophan

F : phenylalanine

Y : tyrosine

L : leucine

V : valine

I : isoleucine

G : glycine

T : thréonine

10 S : serine

5

15

20

E : glutamic acid

D : Aspartic acid

N : asparagine

Q : glutamine

P : proline.

The asterik signs "*" correspond to stop codons (i.e. TAA, TAG and TGA).

Starting above the first line of the DNA nucleotidic sequence of fig. 4 the three reading phases are respectively marked "1", "2", "3", on the left handside of the drawing. The same relative presentation of the three theoritical reading phases is then used all over the successives lines of the LAV nucleotidic sequence.

figs. 2 and 3 provide a diagrammatized representation of the lengths of the successive open reading
frames corresponding to the successive reading phases
(also referred to by numbers "1", "2" and "3" appearing in
the left handside part of fig. 2). The relative positions
of these open reading frames (ORF) with respect to the
nucleotidic structure of the LAV genome is referred to by
the scale of numbers representative of the respective
positions of the corresponding nucleotides in the DNA
sequence. The vertical bars correspond to the positions of
the corresponding stop codons.

1) The "gag gene" (or ORF-gag)

The "gag gene" codes for core proteins.

particularly it appears that a genomic fragment (ORF-gag) thought to code for the core antigens including the p25, p18 and p13 proteins is located between nucleotidic position 236 (starting with 5' CTA 6C6 GAG 3') and nucleotidic position 1759 (ending by CTCG TCA CAA 3'). The structure of the peptides or proteins encoded by parts of said ORF is deemed to be that corresponding to phase 2.

The methionine aminoacid "M" coded by the ATG at position 260-262 is the probable initiation methionine of the gag protein precursor. The end of ORF-gag and accordingly of gag protein appears to be located at position 1759.

The beginning of p25 protein, thought to start by a P-I-V-Q-N-I-Q-G-Q-M-V-H ... aminoacid sequence is thought to be coded for by the nucleotidic sequence CCTATA..., starting at position 656.

Hydrophilic peptides in the gag open reading frame are identified hereafter. They are defined starting from aminoacid 1 = Met (M) coded by the ATG starting from 260-2 in the LAV DNA sequence.

Those hydrophilic peptides are 12-32 aminoacids inclusive

	12-35	######################################	THETOSTAG
	37-46	•	•
	49-79	•	•
25	88-153	•	•
	156-165	•	•
	178-188	•	•
	200-220	•	•
	226-234	•	•
30	239-264	•	•
	288-331	•	•
	352-361	•	
	377-390	•	•
	399-432	•	•
35	437-484	•	•
	492-498	•	•

The invention also relates to any combination of these peptides.

2) The "pol gene" (or ORF-pol)

Figs. 4-12 also show that the DNA fragments extending from nucleotidic position 1555 (starting with 5'TTT TTT3' to nucleotidic position 5086 is thought to correspond to the pol gene. The polypeptidic structure of the corresponding polypeptides is deemed to be that corresponding to phase 1. It stops at position 4583 (end by 5'G GAT GAG GAT 3').

These genes are thought to code for the virus polymerase or reverse transcriptase.

3) The envelope gene (or ORF-env)

10

25

30

The DNA sequence thought to code for envelope proteins is thought to extend from nucleotidic position 5670 (starting with 5'AAA GAG GAG A...3') up to nucleotidic position 8132 (ending byA ACT AAA GAA 3'). Polypeptidic structures of sequences of the envelope protein correspond to those read according to the 'phase 3' reading phase.

The start of env transcription is thought to be at the level of th ATG codon at positions 5691-5693.

Additional feature of the envelope protein coded by the env genes appear on figs. 13-18. These are to be considered as paired figs. 13 and 14 : 15 and 16 : 17 and 18 respectively.

It is to be mentioned that because of format difficulties.

Fig. 14 overlaps to some extent with fig. 13.

Fig. 16 overlaps to some extent with fig. 15.

Fig. 18 overlaps to some extent with fig. 17.

Thus for instance figs. 13 and 14 must be considered together. Particularly the sequence shown on the first line on the top of fig. 13 overlaps with the sequence shown on the first line on the top of fig. 14. In other words the starting of the reading of the successive

sequences of the env gene as represented in figs. 13-18 involves first reading the first line at the top of fig. 13 then proceeding further with the first line of fig. 14. One then returns to the beginning of the second line of fig. 13, then again further proceed with the reading of the second line of page 14, etc... The same observations then apply to the reading of the paired figs. 15 and 16, and paired figs. 17 and 16, respectively.

The locations of neutralizing epitopes are further 18 apparent in figs. 13-18, reference is more particularly made to the boxed groups of three letters included in the aminoscid sequences of the envelope proteins (reading phase 3) which can be designated generally by the formula N-X-S or N-X-T, wherein X is any other possible aminoacid. Thus the initial protein product of the env gene in a glycoprotein of molecular weight in excess of 91,000. These groups are deemed to generally carry glycosylated groups. These N-X-S and N-X-T groups with attached glycosylated groups form together hydrophylic regions of the protein 28 and are deemed to be located at the periphery of and to be exposed outwardly with respect to the normal conformation of the proteins. Consequently they are considered as being epitopes which can efficiently be brought into play in vaccine compositions.

The invention thus concerns with more particularity peptide sequences included in the env-proteins and excizable therefrom (or having the same aminoacid structure), having sizes not exceeding 200 aminoacids.

Preferred peptides of this invention (referred to hereafter as a, b, c, d, e, f) are deemed to correspond to those encoded by the nucleotide sequences which extend respectively between the following positions:

- a) from about 8095 to about 8200
- b) " 6260 " 6310 \
- 35 c) " " 6390 " " 6440 ~

25

d) " " 6485 " " 6620 /

et " 6860 " 6930 -

f) " 7535 " 7630 ~

Other hydrophilic peptides in the env open reading frame are identified hereafter, they are defined starting from

aminoacid 1 = lysine (K) coded by the AAA at position 5670-2 in the LAV DNA sequence.

These hydrophilic peptides are

8-23 aminoacids inclusive

510-523 " 551-577 "

594-603 " 821-630 "

780-603 - "

The invention also relates to any combination of these peptides.

4) The other ORF

20

25

The invention further concerns DNA sequences which provide open reading frames defined as ORF-Q, ORF-R and as "1". "2", "3", "4", "5", the relative position of which appears more particularly in figs. 2 and 3.

These ORFs have the following locations :

ORF-1	•	1	-	5029	•	5316
ORF-2	•	2	•	5273	•	5515
ORF-3	•	1	•	5383	•	5616
ORF-4	. •	2	•	5519	•	5773
005-5	•	•	•	7986	•	8279

The LTR (long terminal repeats) can be defined as lying between position 8560 and position 160 (end extending over position 9097/1). As a matter of fact the end of the genome is at 9097 and, because of the LTR structure of the retrovirus, links up with the beginning of the sequence :

Hind III CTCAATAAAGCTTGCCTTG 9097

15

30

5

10

The invention concerns more particularly all the DNA fragments which have been more specifically referred to hereabove and which correspond to open reading frames. It will be understood that the man skilled in the art will 20 be able to obtain them all, for instance by cleaving an entire DNA corresponding to the complete genome of a LAV species, such as by cleavage by a partial or complete digestion thereof with a suitable restriction enzyme and by the subsequent recovery of the relevant fragments. The different DNAs disclosed in the earlier mentioned British Application can be resorted to also as a source of suitable fragments. The techniques disclosed hereabove for the isolation of the fragments which were then included in the plasmids referred to hereabove and which were then used for the DNA sequencing can be used.

Of course other methods can be used. Some of them have been examplified in the earlier British Application. reference is for instance made to the following methods.

a) DNA can be transfected into mammalian cells . 35 with appropriate selection markers by a variety of techniques, calcium phosphate precipitation, polyethylane

glycol, protoplast-fusion, etc..

5

10

15

20

25

30

- b) DNA fragments corresponding to genes can be closed into expression vectors for \underline{E} . \underline{coli} , yeast- or mammalian cells and the resultant proteins purified.
- c) The provival DNA can be "shot-gunned" (fragmented) into procaryotic expression vectors to generate fusion polypeptides. Recombinant producing antigenically competent fusion proteins can be identified by simply screening the recombinants with antibodies against LAV antigens.

The invention also relates more specifically to cloned probes which can be made starting from any DNA fragment according to this invention, thus to recombinant DNAs containing such fragments, particularly any plasmida amplifiable in procaryotic or eucaryotic cells and carrying said fragments.

Using the cloned DNA fragments as a molecular hybridization probe - either by marking with radionucleotides or with fluorescent reagents - LAV virion RNA may be detected directly in the blood, body fluids and blood products (e.g. of the antihemophylic factors such as Factor VIII concentrates) and vaccines, i.e. hepatitis 8 vaccine. It has already been shown that whole virus can be detected in culture supernatants of LAV producing cells. A suitable method for achieving that detection comprises immobilizing virus onto said a support e.g. nitrocellulose filters, etc., disrupting the virion and hybridizing with (radiolabelled or "cold" labelled fluorescent- or enzyme-labelled) probes. Such an approach has already been developed for Hepatitis B virus in peripheral blood (according to SCOTTO J. et al. Hepatology (1983), 3, 379-384).

Prohes according to the invention can also be used for rapid screening of genomic DNA derived from the tissue of patients with LAV related symptoms, to see if the proviral DNA or RNA is present in host tissue and other

t1ssues.

15

20

25

30

35

A method which can be used for such screening comprise the following steps: extraction of DNA from tissue, restriction enzyme cleavage of said DNA, electrophoresis of the fragments and Southern blotting of genomic DNA from tissues, subsequent hybridization with labelled cloned LAV provival DNA. Hybridization in situ can also be used.

Lymphatic fluids and tissues and other non-lymphatic tissues of humans, primates and other mammalian species can also be screened to see if other evolutionnary related retrovirus exist. The methods referred to hereabove can be used, although hybridization and washings would be done under non stringent conditions.

The DNA according to the invention can be used also for achieving the expression of LAV viral antigens for diagnostic purposes.

The invention also relates to the polypeptides themselves which can be expressed by the different DNAs of the inventions, particularly by the ORFs or fragments thereof, in appropriate hosts, particularly procaryotic or eucaryotic hosts, after transformation thereof with a suitable vector previously modified by the corresponding DNAs.

These polypeptides can be used as diagnostic particularly for the detection of antibodies in biological media, particularly in sera or tissues of persons afflicted with pre-AIDS or AIDS, or simply antibodies in the absence of any apparent carrying disorders. Conversely the different peptides according to this invention can be used themselves for the production preferably monoclonal antibodies specific of antibodies, of the different peptides respectively. For the production secreting said monoclonal antibodies hybridomas conventional production and screening methods are used. These monoclonal antibodies, which themselves are part of

the invention then provide very useful tools for the identification and even determination of relative proportions of the different polypeptides or proteins in biological samples, particularly human samples containing 5 LAV or related viruses.

Thus all of the above peptides can be used in diagnostics as sources of immunogens or antigens free of viral particles, produced using non-permissive systems, and thus of little or no biohazard risk,

The invention further relates to the hosts (procaryotic or eucaryotic cells) which are transformed by the above mentioned recombinants and which are capable of expressing said DNA fragments.

Finally it also relates to vaccine compositions whose active principle is to be constituted by any of the expressed antigens, i.e. whole antigens, fusion polypeptides or oligopeptides in association with a suitable pharmaceutical or physiologically acceptable carrier.

Preferably the active principles to be considered in that field consist of the peptides containing less than 250 aminoacid units, preferably less than 150 as deducible for the complete genomas of LAV, and even more preferably those peptides which contain one or more groups selected from N-X-S and N-X-T as defined above. Preferred peptides 25 for use in the production of vaccinating principles are peptides (a) to (f) as defined above. By way of example having no limitative character, there may be mentioned that suitable dosages of the vaccine compositions are those which enable administration the to host. particularly human host ranging from 10 to 500 micrograms per kg, for instance 50 to 100 micrograms per kg.

For the purpose of clarity figs. 19 to 26 are added. reference may be made thereto in case of difficulties of reading blurred parts of figs. 4 to 12.

Needless to say that figs. 19-25 are merely a reiteration of the whole DNA sequence of the LAV genoma.

Finally the invention also concerns vectors for the transformation of eucaryotic cells of human origin, particularly lymphocytes, the polymerases of which are capable of recognizing the LTRs of LAV. Particularly said vectors are characterized by the presence of a LAV LTR therein, said LTR being then active as a promoter enabling the efficient transcription and translation in a suitable host of the above defined, of a DNA insert coding for a determined protein placed under its controls.

Needless to say that the invention extends to all variants of genomes and corresponding DNA fragments (ORFs) having substantially equivalent properties, all of said genomes belonging to retroviruses which can be considered as equivalents of LAV.

CLAIMS

- 1. A DNA fragment of LAV extending from nucleotide position 236 to nucleotide position 1759.
- 2. A DNA fragment of LAV extending from nucleotide position 1555 to nucleotide position 5086.
 - 3. A DNA fragment of LAV extending from nucleotide position 5670 to nucleotide position 8132.
 - 4. A vector containing a DNA fragment according to any of claims 1 to 3.
- 5. Peptide corresponding to any of those encoded by the nucleotide sequences which extend respectively between the following positions:
 - a) from about 6095 to about 6200.
 - b) " 6250 " 6310-
- 15 c) " " 6390 " " 6440
 - d) " 6485 " 6520
- 8. Peptide characterized by a sequence of amino20 acids deducible from LAV DNA the terminal aminoacids of which extend between the following positions with respect to the lysine (position 1) coded by the AAA at position 5870-5672 in the LAV DNA.

8-23 aminoacids inclusive

25 63-78 "

82-90 "

97-123

127-183

197-201 " .

30 239-294 "

300-327

334-361

397-424

466-500 "

35 510-523 "

551-577 "

or any combination of these peptides.

7. Peptide corresponding to the aminoacid sequences deducible from LAV DNA and the terminal aminoacids of which are positionned at the positions the hereafter counted from the Met at position 1 coded by the ATG sequence at nucleotide positions 260-2:

12-32 aminoacids inclusive

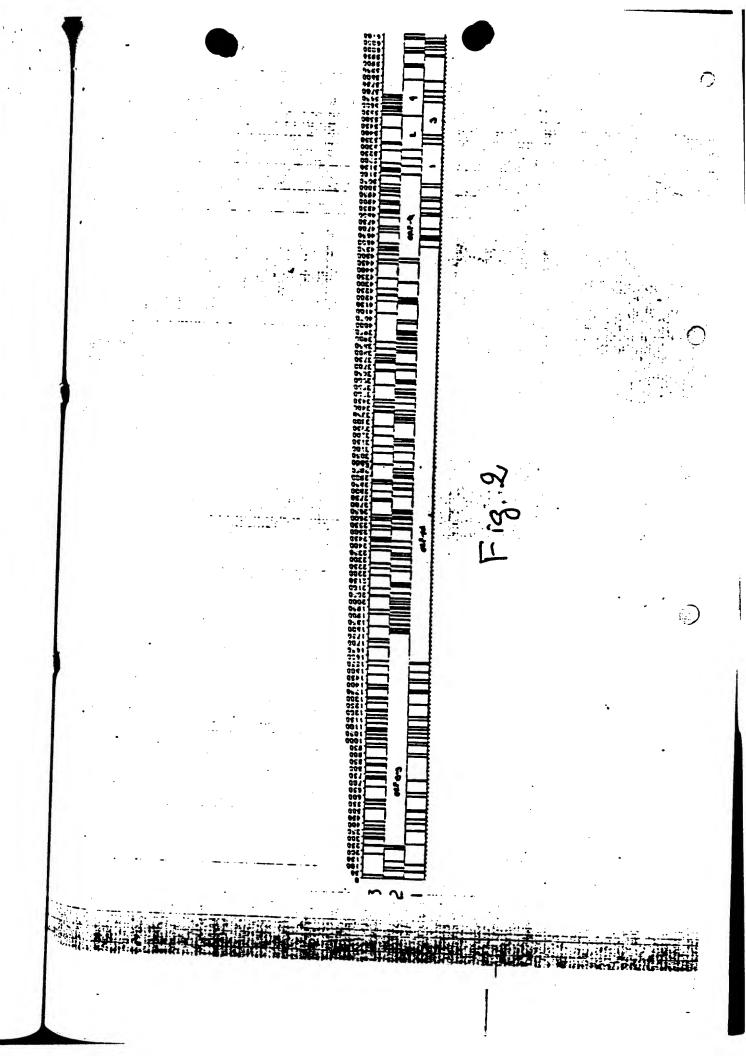
	37-46	•	•
	49-79	•	•
15	88-153	•	•
	158-165	•	•
	178-188	•	•
	200-220	•	•
	226-234	•	•
20	239-264	•	•
	288-331	•	•
	352-361	•	
	377-390	•	•
	399-432	•	•
25	437-484	•	•
	482-498	•	•

and combination of said peptides.

- 8. Diagnostic means containing any of the DNA fragments of any of claims 1 to 3.
- 9. Diagnostic means containing any of the peptides of any of claims 4 to 8.
 - 10. Vaccine compositions containing any of the peptides according to any of claims 4 to 6 in association with a pharmaceutical vehicle.

35

End of transmission



AND THE PROPERTY OF THE PROPER

Fis 3

C. C		01
C. C	* * * 6 * * * * * * * * * * * * * * * *	012 012 012 012 013 014 015 015 015 015 015 015 015 015 015 015
	# E # D C A K A S & L S C C F L A C & L S C C C C C C C C C C C C C C C C C C	S
	C K J G P R T 1 P C P S J P V R H 1 S G L L E T S G L L E T S G L L E T S G L L E T S G L L E T S G L L E T S G L L E T S G L L E T S G L L E T S G L L E T S G L L E T S G L L E T S G L L E T S G L L E T S G L L E T S G L L L L L L L L L L L L L L L L L L	1
		4 D 1 A 5 S F M O D M C A A W 3 W 9 F W W W W W W W W W W W W W W W W W
A C C A A A A A A A A A A A A A A A A A		
1	A	P
TOP	2 H & 6	A 5 S A C A V C T R P D L R T R C S V T P P L 3 N R P R C 3 N R P R C 3 N R P R C 3 N R P R C 3 N R P R C 3 N R P R C 3 N R P R C 3 N R P R C 3 N R P R C 3 N R P R C 3 N R P R C 3 N R P R C 3 N R P R P R P R P R P R P R P R P R P R
I C C C C C C C C C C C C C C C C C C C	•	
	_ <	

- start off. fel # * · _~_; C A D 1 1 V L E N S L F C F A N F N N 1 C F I K V L S K P O S N I R F S P N S C A N S I R R S P N S P Par t anys z X is st L 1 B 2

Z

> ٠ 14.100 LAT

2

2

COPTO SUBJECT SOUR STATES OF STATES

و

×

•

• * _ * •

-

٠ . 9 .

• ų, ٥ £ J •

_ ¥ S K I G P E 13 P V 1 T P V F A I W 4 F D S T K 4 P K L V D F F L V B P T O D F F L L A R T V L N G F N P P T S F C P T P P F L L S R L K T S A F C P T P P F R T S F C L L T S R L L L S R L K T S R L L S R L L S R L L L S R L L L S R L L L S R L L L S R L L L S R L L L S R L L L S R L L L S R L L S R L L L S R L L L S R L L L S R L L S R L L S R L L S R L L S R L L S R L L S R L L S R L L L S R L S R

2

ī

Z

Z

ı

c

.

.

*

-

*

J

>

-

۷

_

S

Ž

- 4

Z

S

Z

Ł

Z

2.3

•• \$\text{c} \text{R} \text{c} \text{R} \text{A} \text{C} \text{R} \text{C} \text{R} \t

~ ٦

E

石ンサ

0 • 6 > _ • _ J 4 ~ u u • × = • • = £ _ • > 4 , •

•

¥

Ŧ • 4 c

Z

Z

₹¢ - 100

20

o

0

Ø

a a	SOSO SUBSTITUTE STATES SUBSTITUTE STATES SUBSTITUTE SUB
ž"	H I V E F V E D T & A C V E A I I B I L O O L L. F I H F P I C C N D I A F O A L L S N K L M C I L C J F W K P O O F F C C L S I S E L G V D I A F O A L L V L O M L M G V L C R S E S H M K M S A I I A V V P F O M W V S I O O M R R V S CAINCRAFAARTRAGGGAAGGGCGACGACGACGAAAACGCGAAAACGCCTHATCAAAAAAAAAA
	0 0 R R A R N G A S P S • T A A L E A S F R S A • N C L V N L L • P V L L S L N R G E Q E S F V D P R L F P M K H P S S O F R T A C F T C V C A R C C P N C 1 E E S R W W S O • 1 L D • 5 P G S 1 O E V S L R L V P L A T V W S V A F I A 1 2 E S R W W S O • 1 L D • 5 P G S 1 O E V S L R L V P L A T V W S V A F I A 1 2 E S R W W S O • 1 L D • 5 P G S 1 O E V S L R L V P L A T V W S V A F I A 1 2 E S R W W S O • 1 L D • 5 P G S 1 O E V S L R L V P L A T V W S V A F I A 1 3
	818 915 915 915 916 916 916 916 916 916 918 918 918 918 918 918 918 918 918 918
3	1 0 1 1 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
3	1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
3	A TO TO THE TOTAL TO A TO THE TOTAL TO THE TOTAL TO THE TOTAL TO THE TOTAL TOT
-	C 2 1 2 1 3 1 4 1 5 1 4 1 5 1 1 1 1 1 1 1 1 1 1 1 1

THE GGG WIPHGS THE GOLGLL WELL HHE WOLGS TO GOLGG STREET WAS WAS WAS STREET CONTROLLED WITH STREET WAS STREET The Letter of Control . r • u _ • • • • . _ _ 4 > -3 ~ ~ -* _ • • 3 -• • Ì 1 ì 3 1 1 3 ⋖ į Ė

V L O E L T S S L L J G T F R N J L S O E R R L G C T G F A R D G L N A S P O T L N I S V F R N L T C L I A R • -I _ ر ه > ... ---• z • u _ 9 • 4 _ • • = u <u>د</u> BAFA ď ď d ح ヹ゚

_